DOI:10.13860/j.cnki.sltj.1986.01.001

从历史发展看中国统计发展方向*

吴建福　(C·F·Jeff Wu)

编者按：吴建福为美国威斯康星大学（Madison）统计系教授，著名统计学家。曾于1985年4—5月来华访问.本文为他在中国科学院系统科学研究所所做报告，本刊独家发表.

这个演讲分四个部份：

一、统计学发展的简史；

二、什么是统计学的发展方向——我们将向哪儿去？

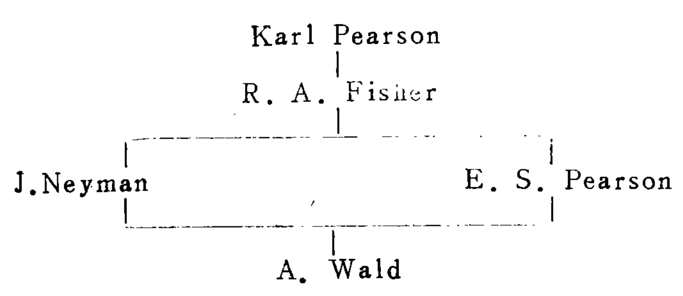
三、对中国统计发展的一些想法；

四、阻碍中国统计发展的几个问题.

我对中国统计学的发展很早就有兴趣.在加州念大学时就知道了许宝騄先生的贡献.后来看了一些学报，知道了一些工作，这几年来往较多，所以有了进一步的印象.

一、统计学发展简史

先画一个表：



这个表不能代表整个统计学的发展历史，只是给出了四个阶段的代表性人物.在20世纪初，Karl Pearson处于统计学的最高地位.他最主要的贡献是给出了χ^2统计量.另一个是所谓Pearson曲线族,后者现在很少有人提了.这些曲线分成很多类，其背景与计算有关，他主要考虑拟合的好坏，而没有重视从资料做统计推断.事实上，他没有推断的概念.他对估计的做法只是矩估计.

R.A.Fisher是一代天才，他的兴趣很广，是统计学中的伟大人物，也是一个伟大的遗传学家.Fisher与K.Pearson争论的焦点是极大似然估计.当然，从现在的观点来看，极大似然估计是很自然的——即将密度函数反过来考虑，于是引进似然函数概念，并使其取极大值从而对参数进行估计.这种方法在当时计算上有困难，但他可以不局限于时代.Fisher最主要的两个工作中,极大似然估计的理论与K.Pearson发生了冲突.但争论的结果是使Fisher成了一位年青（时年30岁）的权威.动摇了Pearson的地位.另一个工作是方差分析，这一点没有与Pearson冲突，很快被接受了.当然，这也不说明Fisher比Pearson强，只是因为Pearson受到了历史的局限.为了解决计算问题，Fisher也做了不少工作，小样本理论从Student开始，到Fisher时代已发展较完整.他也提出资料变换等工具来解决非正态

* 本文由程翰生、冯士雍根据报告录音整理，并经作者审阅.

数据的处理方法。

Fisher提出的似然函数也与假设检验有关联．但数学上并不严格，这个工作是由J.Neyman和E.S.Pearson（K.Pearson之子）做的．Neyman来自波兰，数学很好．E.S.Pearson受到父亲的影响．他们两人合作提出了两个概念：(i)对立假设；(ii)检验的功效(Power)．这些工作中，所用的似然比的概念并不是新的．重要是他们提出了一个统计学的基本定理．Neyman-Pearson引理．现在看来这个引理的证明是初等的，但难在如何提出这个思想．建议大家看两本书，都是有关统计发展的，很有意思：

（1）C.Reid, Neyman from Life,Springer.

（2）J.F.Box, R.A.Fisher, A Scientist,Wiley.

请注意：第二本书题目中用的是A Scientist（一个科学家）。而不是A Statistician（一个统计学家）．这代表了R.A.Fisher对统计的看法．他认为统计是一门科学．而不是一门数学；将统计归入数学是错误的．

我个人认为Neyman的主要贡献是假设检验以及从此而引出的置信区间的概念，从而有了区间估计．虽然Fisher也提出过区间估计．但他的概念不严格，Neyman将概率带进统计，这是一个很大的贡献，他之所以能做到这一点与他的个人经历有关，Neyman的数学很好，在帝俄时代的Karkov受教育，又得到波兰华沙学派的熏陶，后来他来到巴黎又受到Borel的影响．Neyman与Fisher做统计的方式很不一样．Fisher思想很活跃，很多统计思想是属于他的．但是他不喜欢假设检验，这是他的偏见．他也有错误，例如他提出的信任推断（Fiducial inference)现在很少提了．这两个人的争论主要是思想方法方面的．Neyman要求严格，而Fisher虽然提出不少好思想，但较粗糙，因此容易被Neyman挑毛病．

数理统计的严格科学基础是由以上几位大家奠定的，下面要提到的一位是A.Wald．他有两个主要贡献，判决理论（Decision Theory）和序贯分析（Sequential Analysis）．在他之后统计学理论还没有发生如此重大的突破．判决理论是Neyman—Pearson引理的延伸．它对统计方法，例如假设检验、点估计、Multiple Decision提出了一个理论结构以便比较优劣，找出最优法．对商业以及政策决定也有一定影响．序贯分析与实际有密切关系，这是Wald在二次大战期间提出来的，来自于军需生产的需要．这说明强大的需要往往会激发好的理论。军需生产需要实验设计，要节约时间．他的思想突破点是不固定样本大小，将样本大小看成随机的，只要资料中证据足够，就可以停止试验．我觉得思想的突破超过一切．序贯分析的一些问题也引出重要的概率论问题（例如Stopping time）．事实上，概率论的某些方法是统计学家的贡献．

在应用方面，Neyman对抽样也有贡献．但我最喜欢提的是实验设计，这也是Fisher提出来的．他最初在Rothamsted农业试验站工作．在那里提出了方差分析，实验设计的很多思想与数学没有什么关系．农业试验的变差很大，Fisher提出用随机化的概念来解决这个问题，一个实验设计相对于一个模型，但模型很容易出错，而且每一个模型都是近似的，随机化的好处是实际结果不是依赖于模型，这种思想不可能完全用数学解释，他的方法很类似于今天的电脑模拟，但当时没有电脑，只能用不同的设计来看实际指标的变化，实验设计中另一个重要的方法是区组设计，区组设计很像抽样中的分层抽样，引进区组（Block）是为控制变差，因为邻近的的试验结果比较接近．因此将其作为一个区组．另一个观念是因子试验（Factorial design）这主要是Fisher和F.Yates做的．Yates是Fisher的助手，因

子试验的很多工作是他做的．因子试验的思想提得很好，实际中收效也很大，因此，在很短的时间内理论就发展得较为完善．十年内大部份工作就做完了，这也反映了科学发展与实际需要的关系．

以上是统计学一段历史的发展．

二、什么是统计学的发展方向——我们将向哪儿去？

现在有很多问题理论上难以解答。现象太复杂了。因此要考虑是否需要修正现有的理论或者说需要提出一种全新的思想——它基本上离开了Fisher和Neyman的传统思想．这方面的尝试之一是所谓探索性数据分析（E.D.A——Exploratory Data Analysis）．很多人不喜欢E.D.A．但我要指出，这种方法是从搞其它科学（例如，生物学）的人那儿学来的．他们的作法与搞统计学的人不同，不是先把数据压缩成充分统计量之类的东西．但是这种方法还没有形成一门科学，科学要有一般性和逻辑性的．EDA不是这样，不过它所想要解决的问题是存在的．下面介绍一下目前统计发展的趋向，共分为六个方面，问题的选择反映了一定程度的主观性．

（1）非正态多元分析．这里指指标的分布与正态差别非常远的情形．其中之一是投影追踪方法（PP—Project pursuit）．由于指标数目太多，又不符合正态假设，若在原来的高维空间中进行诸如非参数回归，密度估计等处理，会有很大困难．于是设法将数据投影到一个低维空间上，在低维空间上拟合一个曲面，然后考虑其在几个方向的投影，再把几个投影凑在一起．这一思想是J.B.Kruskal提出来的，其中有些理论问题可做．它的主要困难是计算量太大．另一个问题是分类（Classification）．例如在医学中要判断一个病人病情是否严重到需要做手术，由于给出的资料很多，往往难以由一个指标来做判断．这时数据已完全脱离了正态假设．电机工程中的图象识别，心理学中对某种能力的评分也都是类似的问题．对这类问题也可以把高维空间上的数据资料投影到低维空间上，然后排出顺序．还有一个问题是离散数据的平滑化．在实际问题中倒不一定要找一个最优的平滑，而是要找一个较好的平滑．往往我们用一个简单的模型可以做出很好的工作，复杂的模型却得不到好的结果．不过随着数学工具的增加和思想的丰富，有可能解决更多的复杂问题．比如，高维空间的计算是很困难的．拿一个高维空间的样条函数（Splines)计算时可以有很多花样．这个问题在中国也做了不少工作．但没有人将其与统计结合在一起．这即所谓平滑样条，由于观察到的资料有误差，可以用样条来平滑，这是一个统计问题．威斯康星的G.Wahba这方面的工作做得很好．上面提到的问题与E.D.A不同．它们的提法都是很清楚的，而E.D.A则较模糊．

（2）Computer—intensive inference．（即resampling—再取样）．这种方法在电脑方向需求量很大，但是可以解决较多的问题．相对于现有的方法讲，效果是比较好的．本来需要化很多时间来寻找大样本的性质，现在则不需要了．当然，理论工作者还要从这种再取样的方法中找出每种情况下的理论性质．实际工作者就可以根据理论的结果来做自己的问题．利用较多的电脑时间，减少复杂的计算从而避免人工所犯的错误．用这种方法虽然要用较多电脑时间，但大多数实际问题只要中型电脑即可．

（3）小样本的渐近理论．这是一个理论问题．现在大样本的渐近理论已有较大的发展，方法工具已相当完整了．但实际问题中样本往往不太大．微分几何作为一种工具可能对发展小样本理论有用．有时在大样本中计算到第二阶时公式非常复杂，以致于看不出它的

统计意义。由于很多统计定义有度量（metric），可以用微分几何加以解释。这就是说，我们可以用现有的强有力的数学工具来解释一些现象建立一些模型。这方面有成功的先例。我的一个同事D.Bates的博士论文即有此类工作。他发现非线性回归模型中，参数化时可以考虑两类：一类与参数无关，另一类是参数化的。即同一个模型用不同的参数去参数化时得到的区域不相同。这对实验设计很重要，可以找一种参数化方法使参数化的区域很小。我想象统计与微分几何的关系会越来越密切。这一定会对第二阶的大样本理论有所帮助。与此有关的另一个方向是鞍点逼近，（Saddle point approximation）这是个很好的方向。我对一般的大样本理论不是很满意的，这是因为有时大样本理论不能很好地解释小样本现象。

（4）统计计算.这是一个很大的领域。对于统计学来讲，一个方法不能计算等于无用.一个例子是图象处理（image processing）.这个问题和许多其它问题一样本来是属于统计学的，我们却让给了人家，以致于统计学的范围越来越小。现在运筹学的很多分支本来也是属于统计学的.因此，我常讲我们这些统计学家很“笨”，不断地把自己的领地拱手让人.图象本身属于统计范围。因为我们看到的图有随机误差，图象不是很清楚。可以把真正的图象看成一个未知参数，我们看到的形象则是未知参数与一些误差的合成。我们要做的是把误差过滤掉.这是一个工程上的问题也是一个统计问题.美国有几个统计出身的人做了这方面的工作，他们的电脑也是很强的。最优化，应用数学也较好。另一个例子是图表示（graphics），即如何用图形表示统计资料.目前这方面的软件发展较多.理论研究尚少.我所说的理论不一定是数学理论，而是一种逻辑体系。另一个问题是并行算法（parallel computing）；即超大型电脑，它要求硬件有新的发展，也要有新的算法。统计计算要涉及大型计算就要涉及并行算法。这方面统计学家做的工作很少。这与我们受的教育有关。我们学的数学工具中概率是最多的，我想21世纪的统计学家不是这样的。他们要搞最优化、组合数学、离散数学等。因为很多统计与计算有关，这就要涉及到数学的许多新领域。我所在的学校中要求学生除统计之外学一些副科，很多学生选学了电脑，就需要学习这些数学理论。

（5）非随机数据（nonrandom data）。问题的提法是：在取得资料时没有随机化．有时无法随机化。通常称对这种资料的分析为观测研究（分析）（Observational study）。但一个严重的问题是如何做统计推断。以前Cochran曾考虑过这一问题。现在在这方面工作较好的是他的一个得意门生D.Rubin。目前这个领域还没什么数学工作，可是思想很深刻。提出了一些新准则。在中国（社会经济方面）可能也有这方面的例子。例如在计量经济学方面。由于资料来源不是随机的，若做分析（如回归分析）则偏差可能很大。

（6）实验设计与质量改进（Quality improvement—QI）。在威斯康星不讲质量控制（QC），而讲QI。因为要做的是改进而不只是控制。这方面日本人做得最多，中国也有人做。有关的方法在实用上很成功，且还有改进的余地。但是它的理论没有人好好地研究过。我们那里有几个人做这方面的工作。不过目前还没有看到有思想上的突破。有些方法从应用上讲很重要，但不能用现有的理论给出满意的解释。要寻找满意的解释，可能就会提出新的理论与问题。这是一个机会。

三、对中国统计学发展的一些想法

统计学的研究课题有两种选取方法。具体采用哪一种与每个人的实际情况，即背景、能力、经验和兴趣有关。这两种方法是：从理论角度出发和做实际的题目。至于具体选择哪个

课题取决于三个原则：（i）当地（所在国家或地区）的实力；（ii）当地实际应用情况；（iii）目前的世界潮流.

这次来华，有一点体会很深，即中国的统计实力比我想象的强很多．而国外对此了解甚少．这可能是中国人太谦虚了．我听了一些硕士生的工作，工作很好．从数学技巧上讲，有些在美国有可能达到博士论文的水平．中国的工作从技巧上讲是相当强的．至于新的题目，由于与外界脱节而不太好做．你们的工作可以与很多国家相比，例如荷兰、德国．而北欧国家由于理论和应用方面都有工作，情况与你们不尽相同．印度这几年主要做技术性的工作，我看过一些博士工作，中国的程度不低于他们．此外，由于中国前一阶段长期动荡，起点较低，能做到目前这种程度，表明还是很有希望的．问题是你们“广告”做得太少了，要多宣传一些自己的工作．有些成果可以写成英文摘要，也可写些综合报告，以便使外界多有所了解．中国有很多人才，不能总让外面的人来告诉你们要做些什么．

在应用方面中国是最可惜的．中国是一个大国，问题很多．这与一个国家经济发达与否无关，因为即使经济落后，问题仍然存在．其中很多问题需要做统计分析．当然这样做是有困难的，其中之一是对外界统计了解不深，不知道统计学可以解决什么问题．我有一个想法，中国是否可以不用“统计学”这个名词．在汉语中“统计”一词从字面上讲是“笼统地计算”的意思．但是，这并不是各位心目中的统计学．我们是否可称之为数据科学（Data science）．这个名称比较真实地反映我们的工作．而现在人们认为统计就是统计局所做的哪些工作．他们的工作主要是收集数据，并加以初步整理．这些倒是符合“统计”一词的，但这只是统计学的一部份内容．中国要想很快地赶上、超越别人，有一个捷径，就是先在名词方面创新（笑声）．这样，我们向别人自我介绍时，也不自称为统计学家（statistician），而称为数据科学家（Data scientist）．台湾其实也有类似的问题，一提起统计学家总使人想起会计（笑声）．听说中国教育部已批准在一些院校成立统计系，也有的叫数理统计系，不如丢掉“统计”二字而称之为数据科学系．中国科学院若有机会成立一个所，不妨也起一个类似的名子．名称响亮了经费肯定增加（笑声）．

统计活动有三个阶段，我称之为三部曲：（1）数据资料采集；（2）资料的分析及推断；（3）结果的解释并与问题的结合（interpretation）．第一阶段有两类方法．对于可控的总体（例如工业生产）可采用实验设计的方法．若考察与人和社会有关的问题则采用抽样调查．这时总体难以控制，但可以用诸如分层的手段来选取不同的收集资料方式．取样是非常重要的，不管对资料的分析做得多棒，样取不好结果很难令人信服．第二阶段是搞统计的人搞得最多的，这里就不赘述了．不过要指出的是，在做这步工作时开始可先做些简单分析，不一定一开始就很正规（formal）．第三阶段是我们知道得最少的．也是中国的一个弊病．我们回答科学问题或政策问题不是完全只凭资料，资料只是一种辅助．一个好的统计学家可以把分析结果转化到实际问题中去．但如果对实际问题了解不深就会出错．也就不能达到使统计学为其他科学服务的目的．不过，中国国内这方面工作做得好的人可能不被重视．我认为，假如有人利用统计作为工具解决了一个别人无法解决的问题，他应该是统计界的英雄．当然做这方面工作的人不一定是所谓的统计学家．他可能是心理学家或经济学家．这也说明，我们不仅要培养合格的统计学家，也要教会别的领域的专家使用统计方法．

下面谈一下中国可能发展的方向和领域．

（1）农业．农业统计看来没有什么可做的，其实不然．这方面虽然很多工具都有了．

但每个国家情况不同．中国农业就与外国有很大区别，有自己的结构与特色．我认为实际问题有特色，就可能引导出一些有意思的理论问题．例如，在育种、林业、畜牧等方面各有自己的问题．可是从事这方面工作的人往往没有与统计学家在一起研究，造成脱节．

（2）工业．工业与农业不同，农业是一分努力一分收获，但工业进步往往是飞跃．在这方面，统计学可用的地方很多．中国有不少人做这方面的工作，搞了很多推广，也能看到一些实际问题．但是为什么不能从中找出一些好的理论问题来研究呢？我不相信没有好的问题．可能做这方面工作的人有一个框框，习惯于用现成的方法去套实际问题．当然做比不做强．但是要发展新的方法科学才能进展，要有好奇心．在工业中应用统计学中国有基础，已做了相当的推广工作，有些工厂也很支持．

（3）医学统计．中国和美国不同，美国人少钱多，而中国恰好相反．在美国一种新药在上市之前要经过动物实验，临床实验．要用统计来控制实验结果，以避免假象．要求是很严格的．在美国的生物统计很吃香．这也表明客观需要是很重要的．像生存分析（Survival analysis）Case control study，Cox regression model 等新方法都是从实际需要出发而发展起来的．诸位条件比我好．在美国有很多好题目落不到我手．这里另一个要注意的是，在这方面中国不能照搬美国的作法．美国太浪费．例如癌症研究搞了好久，化了很多钱，效果不大，可能还不如一些简单的预防性工作（如食物、生活习惯）．这是美国的缺点．你们不要学．要去芜存青，此外向外国学习统计学，也不要只限于美国，有些国家化钱较少，但也搞得很好，例如英国．中国人很少去英国学统计，这很可惜．

（4）政府统计．政府统计有两个方向——抽样调查和计量经济．计量经济是把统计应用到经济的结果，过去中国对计量经济不重视，现在已开始重视了．其实计量经济学无论对计划经济，市场经济还是中国目前这种混合经济不仅很有用，而且很重要．一般利用计量经济模型预测市场有上千上万个方程式．不过在中国搞这方面工作也不能都照搬美国．你们刚刚开始这方面的工作，这也有有利的一面．最好要有很好的统计学家加入这方面的工作．搞计量经济的统计学家要学经济，你学他、他学你．中国有一个很特别的经济结构．会发现一些教科书上没有的新问题．真正从事这方面工作的人（尤其是年青人）将来希望很大．至于抽样调查，（这方面中国不要学苏联）由于中国的人口与社会结构与其它国家大不相同，不妨大胆采用一些新方法，不要过于保守．采用新方法可以达到节约人力、物力的好效果．当然，这里也有有实际经验和有理论知识的人合作的问题．

除以上讲的之外，还有一些其他的领域，象可靠性工程，如能与搞工程的人结合得好，中国有希望做出一些新东西．

四、阻碍中国统计发展的几个问题

我个人认为有以下几个问题比较严重．

（1）理论与实际没有很好结合．这是最严重的问题．我上面讲过统计有三部曲．你们没有很好地将这几部分联在一起．这里有一个主要原因是，在中国统计工作的好坏要由数学家来判断．数学家不懂统计，用数学的准则来判断统计是很危险的．我希望各单位搞统计的尽早与数学分家．事实上，数学系中统计与别的学科差别最大，从历史上看，从数学中分出来只有统计学和计算机科学二家．

另一原因是在中国做了实际工作——例如咨询——提级时不算数．这是不对的．我有一

个朋友，数学不算最好的，但很受尊重，因为他在农业统计中打出了一条路子。虽然他在数理统计杂志，不太登文章。他的文章大多发表在生物、林业杂志上，但是他的工作影响很大。首先他打开了门路，使农学院与统计系合作。并请了二个助教授参加他的工作。这样等于在农学院有了三对“耳朵”。有好的问题提了出来，学生也有机会到那里实习，受到做实际问题的训练。学生受过这方面训练后本事比较大，不是只会证明定理的。至于对方，他们原有的问题自己不能做，在我们的帮助下就可以解决了。老实讲，这样干是比较难的，要有很高的境界。能两边跨的人才很少。中国大，这方面条件好。尤其北京是首都，在这里什么问题都能遇到。

（2）近亲繁殖。这在中国也是个严重的问题。大家都怕自己最好的学生送给别人，对科学的发展很不利。听说科学院最近有个规定，百分之几十的研究生毕业后要离开科学院，这是个好开始。事实上，你看人未必能看准，每个人的潜力不那么容易被发现。有的学生自己能闯，毕业论文不一定好。但他聪明、独立，有可能闯出新领域。学生受老师的影响太大，难以做新的题目。把最好的学生放走不一定有坏处。而从别处来的学生到你这里，可以带来新工具与思想。可以考虑互换学生。另外看成果不能只看论文多少。要看工作中有多少新想法。搞四化，就要和世界最好的成果比。中国数学的技巧好但由于近亲繁殖，可能不易出新东西。学生到了国外去就可以选择新课题。

（3）人事僵化，本来中国统计界中的强手就不够多.每个单位又都留下最好的，致使每个单位最好的人更少，达不到“临界质量”（critical mass），难以起飞。从很多国家发展科学的模式（包括苏联）来看，当科研人员少时，起先最好能集中在两、三个中心内。在这些中心里，强手数目达到了临界质量。当然有些单位要受些损失，但中心能培养出第一流的学生，分配出去。因此，从长远来讲对全局有利。

（4）冗员太多。这个问题大家都明白，就不多谈了。要提的是，不仅非科研人员，科研人员中也要减少冗员。

（5）文献不足。我认为每个单位一定要有以下几份杂志：

Annals of Statistjcs, Journal of American Statistical Association, Journal of Royal Statistics Societg Ser(B), Biometrika, Technometrics, Biometrics.

此外，还要有下面这种杂志作为查文献的工具：

Current Indenx to Statistics.